European Commission

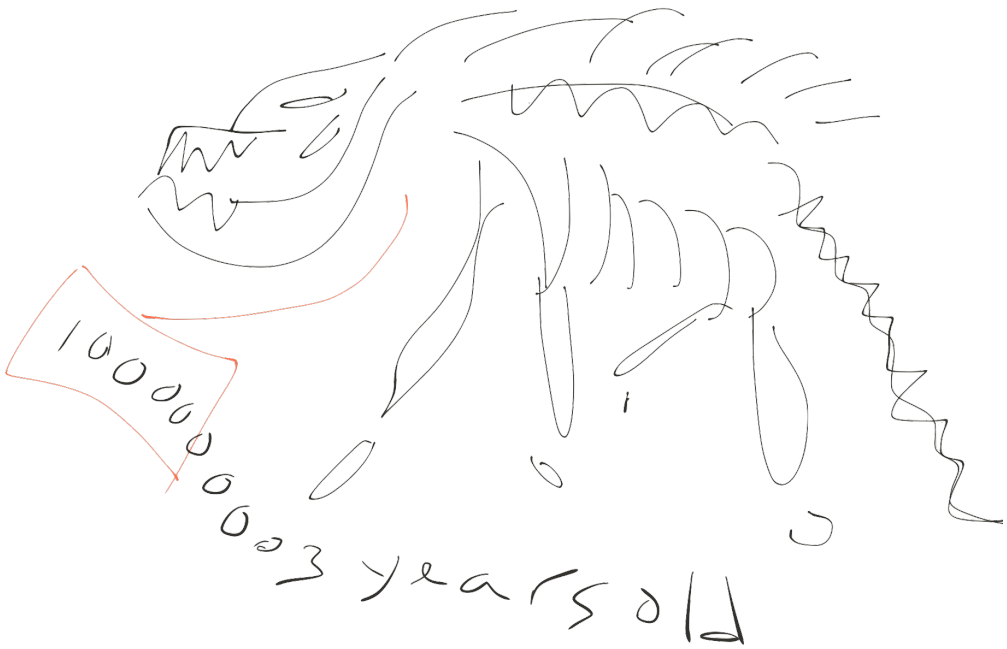
JRC CONFERENCE AND WORKSHOP REPORTS

Significant Digits: Responsible Use of Quantitative Information

Inspirational Workshop 2

Ângela Guimarães Pereira
Jerome Ravetz
Andrea Saltelli

2015



Joint Research Centre

Significant Digits: Responsible Use of Quantitative Information

This publication is a Conference and Workshop report by the Joint Research Centre, the European Commission's in-house science service. It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

JRC Science Hub
https://ec.europa.eu/jrc

JRC98438

ISBN 978-92-79-54261-9 (PDF)

doi: 10.2760/9793 (online)

How to cite: Guimarães Pereira, Â., Ravetz, J., Saltelli, A. 2015. Significant Digits: Responsible Use of Quantitative Information. European Commission. Doi 10.2760/9793.

“The appeal of numbers is especially compelling to bureaucratic officials who lack a mandate of popular election or divine right; scientific objectivity thus provides an answer to a moral demand for impartiality and fairness; is a way of making decisions without seeming to decide.” T. M. Porter, 1995.

Table of Contents

Summary

The DG JRC organised in 9-10 June 2015 a workshop "Significant Digits: Responsible Use of Quantitative Information"[1], mostly targeted at European Commission's colleagues, which invited to reflection on the problem of irresponsible misuse of quantitative information in policy relevant matters. The event was opened by **Vladimír Šucha**, Director General of DG JRC, producing a tide of 'insignificant digits'.

Vladimír **Šucha** welcomed this workshop as a departing point to engage well-known criticisms of the irresponsible use of quantitative information to underpin policy making. He argued that in many fields numbers are needed and therefore we should think of ways to enhance their quality, as well as obtain support from other types of analysis such as qualitative assessments and major integration of social sciences. He also remarked that the request from the policy making spheres for quantitative information is linked to the promises that many scholars make about the value of their numbers, the models that generate them and the data from which they are crafted. Hence, he requested that this workshop be carried out in a constructive manner in order to address the problems we all know exist, providing insight into the changes needed to challenge the current problems of robustness and quality that many policy issues are facing.

Hence, the *Significant Digits* workshop looked at the quality, high and low, of the uses of quantitative information to inform policy making. Through several examples, it looked into uncertainty and the ways in which it is tackled in very important sustainability related policy files; *hypocognition* and "socially constructed ignorance" as symptoms of intentional or naïf assumptions and simplifications of complex issues, namely in the energy, food and agricultural sectors; the confounding of scales of analysis when using quantitative information; serious misconceptions about probability leading to *quantifauxcation*; the lack of social robustness of the indicators and models that are used to deal with complex societal issues; the significance of quantitative information in a plurality of perspectives where different sources of credibility and legitimacy are at stake; farfetched assumptions in predictive models where there can never be adequate knowledge for effective support; and realisation that quantitative information embed narratives and disciplinary perspectives that represent specific perceptions of reality. **The workshop recommended different ways of working reflexively and deliberatively within imperfections, as uncertainty is intrinsic to complex systems; in other words, a strong call for social sciences was put in all affairs where science is relevant.**

This report summarises the contributions of all speakers as well as the discussions held during the workshop. It concluded that the analysis of bad practices is useful but not enough, and taken in isolation it may indeed eventually become counterproductive. Hence, the report concludes with recommendations for a way ahead at the JRC, suggesting activities inside and outside the JRC. Training at different levels of operation at the European Commission (scientific and policy officers alike) is an option to be considered. We suggest that an extensive embedding of the users and trainers of the JRC's reflexivity tools in its regular work would provide a great service to the European policies it supports, and also to the policy process in Europe and worldwide. This work could then be extended outside.

[1] The relevant Science Hub page is: https://ec.europa.eu/jrc/en/event/conference/use-quantitative-information

The workshop contributions

Jerome **Ravetz** offered an historical perspective on how and why the crisis of 'evidence-based' policy come to be. He claimed that its deep origins lie in a certain conception of knowledge, which can besummed up as a *faith in digits as nuggets of truth*. As he pointed out, this is so deep in our modern scientific culture, "its exposure and correction will not be a quick or straightforward process". Taking as example the on-going economic crisis, **Ravetz** questions the faith and reliance on numerical data and mathematical methods of economists and other social practitioners, alike which have showed us how 'everything' can go wrong when they are applied incompetently. **Ravetz** noted that this delusionary faith has complex roots, in traditions of philosophical thought and social practice. Focusing on digits, as a core element and symptom of the pathology, he suggested that "since so many policy issues now involve quantities with 'not even one significant digit', we need an appropriate **new arithmetical language, based on 'soft numbers' using 'sparse digits'** and a dynamical graphical arithmetic of uncertainty and quality."

> "The ideology of 'speaking truth to power' is now obsolete, even antique. But where do we go from here? (…) [An] approach to quality (…) certainly involve[s] an 'extended peer community'. How will the mainstream scientific enterprise respond to such a suggestion? However it is implemented, it would involve a ceding of power and legitimacy to some external agencies. (…) For that, we will need a dialogue, and that dialogue must involve a deeper critique of the dominant practice, so that it is universally seen as unsustainable, both practically and morally. "
> *Jerome Ravetz*

Jeroen **Van der Sluijs** discussed the different types of uncertainty and how they are dealt with at the science-governance interface. He identified three framings of uncertainty that correspond to three different understandings of it: the "deficit view" – uncertainty as a temporary imperfection in our knowledge that will improve with more research; the "evidence evaluation view" which recognises multiple voices in science but mediates this into a shared consensus view and; the "complex systems view/ post-normal view" that sees **uncertainty as intrinsic to complex systems** and requires working **deliberatively** within imperfections. Uncertainty is much more than a number and there are many dimensions (technical (inexactness), methodological (unreliability), epistemological (ignorance) and societal (limited social robustness).

Drawing from several examples, **Van der Sluijs** described different ways of dealing with uncertainties within policy making circles, namely listing: application of Bayesian methods, consensus seeking, nihilism – i.e. dumping the science, the precautionary approach, weighing of experts reputation or trust on experts based on personal views or best fit for the policy agenda and finally what he described as the post-normal approach which seeks to explore the relevance of ignorance and devise ways of dealing with it. Through a detailed discussion of cases of hyper-precise numbers e.g. the claim "*7.9% of species are predicted to become extinct from climate change*" (Urban, 2015) and the evolving

> "We always simplify, so we tend to set very limited system boundaries to keep scientific assessments manageable, but we have to understand the impacts of these design choices on the validity and scope of the conclusions of such assessments. Stakeholder engagement is ever more important. They are a useful resource in co-framing the problem and identifying what is relevant to address, they can also provide useful information and data that scientists have no access to otherwise, and they can also be a critical resource in quality control and extended peer review."
> *Jeroen van der Sluijs*

definitions of the 1.5-4.5 degrees climate sensitivity range he has offered a NUSAP based methodology to underscore the strengths and weaknesses of numbers offered through these concepts, including a close look into their pedigree, concluding that we **need more qualifiers of scientific information** and improvement in terms of **craft skills with numbers.**

> "The damages of socially constructed ignorance are generated by either (or both) endorsement of sloppy quantitative analysis (bad models or indicators); or endorsement of irrelevant storytelling. (…) How many people do really believe that the most relevant problem that humankind has to face now is to prevent a 78cm rise in the sea level in the year 2100? (…) I believe that ignorance is more due to irrelevance of the chosen storytelling than to problems with the quantitative representation of problems" *Mario Giampietro*

On his fight against what Lakoff in 2004 described as "hypocognition" and what Ravetz in 1986 described as "socially constructed ignorance", Mario **Giampietro** outlined the importance of context, scale, storytelling and storyteller.

Through some examples of indicators in the energy and sustainability[2] fields he showed that these issues are often overlooked and not even understood. He argued that we do have a problem with the way quantitative science is used in the field of sustainability right now because it is used with the purposes of prediction and control (risk assessment, optimal solutions, best course of action. etc.). Of course, when dealing with complex systems we can still use quantitative science to gather useful insights. He proposed to move away from assessments based on "a single set of numbers" to assessments based on "expected relations over several sets of numbers" i.e. **quantitative analysis based on patterns and grammars[3] that he called "quantitative story telling"**. This implies a radical revolution in the use of numbers in science for governance. In the Cartesian dream of prediction and control numbers (supposed to be generated by good models developed within a sound story-telling) are used to indicate what is the best thing to do (to individuate optimal solutions). On the contrary in quantitative story-telling numbers are used to check the quality of the pre-analytical narratives determining the quantification and the usefulness of the chosen story-telling determining the relevance of the analysis.

[2] See a critique on the ecological footprint indicator in Giampietro, M., & Saltelli, A. (2014). Footprints to nowhere. *Ecological Indicators*, *46*, 610-621.

[3] A "grammar" is defined as consisting of a taxonomy (defining what is relevant according to predominant perception); a lexicon (choosing what is observed and represented) and a set of production rules establishing causality in the chosen representation, deciding what should be considered as either dependent or independent variables.

"Another fantastic indicator is the "ecological footprint". At the world level the only relevant factor determining an increase in the ecological footprint is the growing requirement of land to capture CO_2. But this requirement is calculated in a very bizarre way. First of all, when dealing with energy security they consider only the sink side: how much you need to capture CO_2, but not how much land you need to produce the energy emitting CO_2. This implies assuming that fossil energy supply will be available forever and that the forest (fixing CO_2 on the same hectare) will grow forever. In fact, if you are assessing the sink capacity of the area (kg/m^2) needed for catching a flow of CO_2 (kg/sec) either the sink capacity has to grow forever (the value of kg that can be stored per m^2 keeps growing in time) or the requirement of hectares of forest must increase in time. That is, unless you have an integral in time you cannot map a flow (kg/sec) on a sink capacity (kg/m^2); it is dimensionally wrong. We can only conclude that one of the most popular analytical tools used right now to study sustainability is based on an equation where the dimensions on the left are not the same as the dimensions on the right. How is it possible? It is important to answer this question because it tells about the systemic problems encountered by the semiotic process used for checking the quality of many sustainability indicators…" *Mario Giampietro*

Through an example from the fisheries sector in Norway, Dorothy **Dankel** illustrated how entrenched positions of different stakeholders determine what needs to be known, what information needs to be published or processed and how those are negotiated by all stakeholders. In addition she showed the symbolic importance of the large number of significant digits that arise from the stock assessment models that support the decision making process, suggesting that **those digits are not about precision of measurement but about the outcome of negotiations as each extra tonne of assessed fish has political value**. Her story of Norwegian pelagic fishermen taking the lead on surveying stocks of pelagic species in collaboration with national authorities but not being able to use the data they collected by decision of the same scientists is a classic example of clashes of legitimacy about knowledge governance. Ultimately, such situation puts in jeopardy trust in reference scientific institutions with *de facto* scientific and political credibility. In addition, she made a clear case for lack of quality assurance when the same scientists engage in a multiplicity of roles as the developers, judgers, reviewers and messengers about stock assessments and quotas. She suggests post-normal framings, namely NUSAP as a way **to work deliberately within imperfections** as numbers are essential for advice in the fisheries case; the excuse "it is the way we do it" to avoid more reflexive procedures is not good enough anymore.

"The pelagic fisherman in Norway, who are well organised and wealthy, agreed with scientists that because of climate change the distribution of pelagic fish has expanded. But because of the expansion the traditional survey cruises budgeted to survey just some areas, cannot absorb the entirety of the distribution, the fishermen suggested that the survey needs to be expanded. (…) as the government would not finance this further, the fishermen agreed to pay for it, putting ½ million EUR in January [2015] to help the scientist to have a better survey coverage. When the survey was over, the lead scientist oublically (in the newspaper) recognised it as the best survey they have ever had; the collaboration had been great; unprecedented coverage of the stock. But then in February, the scientists came out with a new stock estimate number, but keeping the data close to their chest. They said they would not make the survey public, that there would be a relative estimate for the state of the stock therefore not eligible for revising the quota. [So, the fishermen were unhappy and a great media attention was given to this, about which the scientists responded very assertively]. (…) The pelagic industry is so convinced that the way they are doing science is wrong that they will finance the science themselves." *Dorothy Dankel*

"Monte-Carlo simulation process is a non-sense", *John Kay*

John **Kay** challenged the idea of the real model out there. Using two compelling examples of models some use everyday (the London Tube map) and the WebTAG forecast model[4] he has illustrated on one hand the persistent idea that models fit the territory and on the other hand the inability to have the knowledge these models assume. The results often discredit of the whole enterprise of modelling. In his view, in many of these models the future is assumed to be essentially similar to the present, when in the particular model he described with regards to driving patterns, it is hard to have any idea about what the future is like. Also uncertainties within the model are often ignored or dealt with in an unsatisfactory way – for example, how do we know whether a particular concept or variable we use today will be at all relevant in 50 years? Another group of problems arise in economics, for example use of "analogue economies", simplified assumptions such as the behavioural approaches that consider one agent as a representative of all behaviours. All models impose an axiomatic of the world in the world. **Kay** suggested that we follow Lyotard's view that **we need to move away from grand narratives *vs.* little stories, i.e., from grand models to models that are actually useful in particular problems and are not useful for other purposes**. He proposes using models in different ways, i.e., simpler models that identify key factors that influence assessment, deciding at each stage if quantifications are really relevant not adding to confounding arguments. Moreover, he argued that piecemeal assessment of individual components of large projects is more useful than black box analysis for policy making. Finally, he suggested **we "abandon completely the search for standard templates and universal models"**, arguing that the belief that this is possible is a fundamental methodological misconception.

> "If you ask what is wrong with it [predictive models], then I think you need to start with the basic concept from which it begins which is you are doing an exercise essentially like the one NASA was undertaking and you ask yourself suppose we knew everything we might want to know about the system we are undertaking, what would be the information which we would need in order to make that assessment? So, you imagine you know everything you could possibly know about the world and then you could construct a model on that basis; you then encounter the problem that you hardly know any of that information so what do you do? You make it all up. (…) [Several implications, namely] since most of the numbers are invented you can usually select the invented numbers to get whatever results you want. For example, for the project in the UK of constructing a train high speed link between London and Birmingham." *John Kay*

Andrea **Saltelli** illustrated the elements of crisis of trust and quality in the current scientific endeavour, focusing on reproducibility - in both natural and social sciences, and an increasingly dysfunctional peer reviewing processes. As noted by scholars such as Ravetz (1971) forty years ago and more recently Mirowski (2011; 2013), science as a public good has become a commodity that can be described as techno-science. This has lost its original ethos and quality. The present predicaments of science's quality control mechanism cannot be solved 'from within' with technical fixes and better incentives. **Saltelli** called for more

> "What better way to express the crisis in trust than using the words of the Pope? (…) This is a speech given to the Parliament in Strasbourg in November of last year. If you look, there is an aspect of these remarks where, at the beginning, we have lost the great ideals. However, then the loss of trust is associated with the fact that institutions are perceived as aloof and are engaged in laying down rules that are seen as insensitive to individual people, if not downright harmful. The legitimate question is whether some or most of these rules are the result of some kind of misplaced evidence for policy." *Andrea Saltelli*

[4] See e.g. https://www.gov.uk/guidance/transport-analysis-guidance-webtag

attention for the lessons of sociology of science (such as science and technology studies) and for emerging forms of science as possible elements for a way ahead.

Evidence based policy needs to pay attention to existing problems of science. Quoting Lyotard (1979), Funtowicz (2006) and others, he argued against the prevailing demarcation model that separates facts from values and science from policy. After reviewing some egregious cases of error, abuse and misuse in modelling, **Saltelli** proposes a vigilant attitude tot quantifications, including **"sensitivity auditing" as a means to both construct and use models and indicators for use in policy**.

[with regards to the financial crisis], "the scale of analysis that is used to understand what was happening at micro-level does not explain what was happening at the macro-level, the propagation of risk and the loss of information caused by the pooling of many derivatives. On one hand, there is the rational agent (profit maximisation agent) which according to micro-economic theory is supposed to lead to economic growth through the invisible hand of the market and if individual agents are doing Ok this should result in growth for all, but in practice what we saw is that in the long term this led to distribution of risk and systematic loss of information. At the macro-level the story told was differently, what we saw was that year after year the GDP was growing whilst we ended up in an economic crisis, not being completely clear about what happened. So, in this case despite the rich knowledge basis, it didn't provide any quantitative information about what was going on, even if there were a lot of numbers being produced." *Zora Kovacic*

Zora **Kovacic** focused her talk on quantitative evidence in policy issues where there are great uncertainties, offering three contemporary case studies (GDP, water, and the financial crisis of 2007-08), which feature different types of dissonance on narratives, perceptions and representations. She offered multi-scale analysis as a vantage point to discuss some of those dissonances. As she asserted, modelling indicators implies the choice of the field of expertise that informs evidence and the choice of specific disciplines as opposed to others, as well as representation of some narratives to the exclusion of others. She proposed multi-scale analysis is proposed as a tool to understand what knowledge claims tell, to identify the limits of a representation and to identify knowledge gaps when looking at evidence based policy making. **Kovacic** further offered reflexivity as a tool where questions such as: who defines the relevant representation of an issue? To what extent it is worth producing more accurate measurements to give better advice to policy? Is what we are doing really useful? What are the normative components in the technical decisions that are made to produce evidence? Is the problem definition relevant?

"George Box once said, "All models are wrong but some are useful". But before using these models, some questions need to be asked. What makes a model useful? Who defines the purpose and usefulness? One example is GDP, which was conceived during WWII with the objective of measuring productive activities in the economy, to investigate if countries were capable of engaging in the war. But is this model still useful? Does it make sense now, to measure productive capacity in service economies?" *Zora Kovacic*

"How does probability enter a scientific problem? It could be that the underlying physical phenomenon really is random (e.g. radioactive decay, quantum mechanics); it could be that you deliberately introduced randomness (e.g., a deliberately randomised experiment, a random sample); it could be that you use subjective probability to attempt to quantify personal uncertainty; or the probability could exist only within a model invented to describe the phenomenon (metaphorical probability, e.g., claiming that earthquakes occur "as if" a casino game). When the probability comes from a model, one should ask whether the model has been tested and demonstrated to be adequately accurate for the task. For instance, there aren't enough magnitude 8 earthquakes in a particular region to tell whether any model is useful." *Philip Stark*

Whilst there are meaningful ways to use probabilities, Philip **Stark** described a series of serious misconceptions about probability, giving examples from actual applications that underpin policy making. He described the process of "quantifauxcation" whereby meaningless numbers are produced and then treated as if they had meaning merely by virtue of being numerical; he claimed that this is a common source of numbers used as the basis for policy decisions. Cost-benefit analyses often assume that all costs can be put on a common scale, whereas even common sense can assert that some are incommensurable - e.g. quantifying human life in dollars. But multi-dimensional scales cannot always be well ordered. Stark argued against claims that rationality requires quantitative cost-benefit analyses. Rather, if there is no rational basis for the underlying numbers, how can a decision that relies on the numbers be rational? Another example of "quantifauxcation" is many activities under risk analysis: What if probability doesn't apply to the phenomenon in question, if the consequences cannot be reduced to a one-dimensional scale, or if the consequences are not known?

"Quantifauxcation: Assign a meaningless number, then pretend that, since it's quantitative, it's meaningful. [This] usually involves some combination of data, pure invention, ad-hoc models, inappropriate use of statistics and logical gaps." *Philip Stark*

Another example of "quantifauxcation" is what **Stark** described as "Cargo cult confidence intervals", i.e., when the formal calculations of confidence intervals are applied to samples that are not random samples. For instance, it is common to use the mean and standard deviation of an ad hoc collection of models or estimates to make a "confidence interval" or "probability" statement – seen over and over, e.g. in IPCC reports. In order to overcome issues of replicability, he suggested "preproducibility", i.e. a description that includes those things that we may not, with advantage, omit, which is a current problem in science. Without describing the assumptions of the model, methodological procedures, etc. there is just a story, not scientific evidence as its claims cannot be verified.

Discussion and Way ahead

> "Do you need a number to see that it is a bad idea to put a nuclear site close to a major fault?" *Mario Giampietro*

Key Quality Issues

Through this workshop we have identified key quality issues that lead to, or result from, bad practice and irresponsible use of digits in the scientific practices associated with policy making.

1. **Insignificant digits.** The workshop exposed a great number of examples that show that there is indeed a serious problem. This becomes apparent in a variety of contexts: indicators and models' assumptions; rhetoric concealed in the apparently bare number; the fabrication of numbers in order to run otherwise un-runnable models; indicators that are affected by type 3 error, i.e. *indicating* irrelevant or confounding state of affairs; or phenomena described at the workshop as "quantifauxcation", "hypocognition". The importance of knowing who is generating those numbers and deciding what needs to be measured, for what purpose and context, as well as the world-views that are sustaining such numbers, calls for urgent societal enquiry involving a broad range of stakeholders.

2. The **importance of the story telling**: the workshop emphasised the importance of the story told and of the story tellers behind the numbers that are generated and used in policy making. It emerged that there is a need for clarity about the implicit story behind the numbers that are generated and used in policy making. Those background stories shape the assumptions behind the models and statistical techniques, as well as in the intended rhetorical effects of apparently bare numbers.

3. A matter of **institutional (dis)trust**. The workshop identified striking examples where institutions' credibility and legitimacy are put at risk because of poor practices in the generation, use and communication, or even the concealment, of quantitative information. These practices put institutional credibility and citizen trust in institutions in jeopardy, adding to the on-going crisis of legitimacy in policy-relevant science affairs.

4. **Discrediting science**. Further to the issue of credibility in institutions that generate and use numbers in the name of science, the discrediting eventually also affects the scientific endeavour as a whole, becoming generalised to realms where digital information could in principle be significant. This tendency is very visible in the way that uncertainties are played out at policy level.

5. In the matters of concern, an isolated individual cannot decide what is societally relevant and therefore what needs to be counted. The workshop emphasized quite strongly the need for **more broadly based societal research** when numbers are involved. What is relevant? What is significant? These are questions that cannot be left to science or policy elites alone; in fact reasoning with numbers for policy without broad social engagement is close to an oxymoron.

6. Throughout the workshop a number of **reflexivity tools** have been offered to ensure the development of **craft skills** with numbers. This would enable the scrutiny of the **quality** (including the significance) of numbers and digits produced to underpin policy and other types of action. Frameworks and tools have been proposed such as: tools that investigate

processes of confounding and hypocognition, including quantitative story telling and sensitivity auditing; working deliberately within imperfections as in post-normal science frameworks and the NUSAP methodology; *preproducibility* to avoid statistical misconceptions and '*quantifauxcation';* ensuring social robustness by making use of more social research namely dialogue and deliberation methods about the matters of concern.

7. Reinforcing **awareness**. Through a variety of examples, the workshop showed standard problems of incompetence that affect the way numbers are produced and subsequently deployed: incompetent application of statistics; de-contextualisation of models and indicators (being used outside of their field of applicability, inappropriate generalisations, compression of scales, etc.); ignoring of epistemic and social uncertainties. The workshop suggested that merely raising awareness of such errors is probably insufficient. It needs to be reinforced not only at the scientific level, but also at institutional level. Quality assurance is a collective endeavour for the whole community including stakeholders, not to be left to specialist groups.

8. **Crisis and reform.** This technically oriented workshop took place in the context of a general and intensifying crisis in the quality and trustworthiness of policy relevant science, and hence in the responsibility of those who conduct and manage it. Before the workshop took place there was an exposure of decades of misconceived advice on nutrition, based on the 'lipid hypothesis' starting with a paradigm-setting but deeply flawed 'seven nations study' of dietary fat and heart disease. Since the workshop the world has learned how scientific studies of diesel engine emissions were deliberately and systematically perverted by one of the world's most reputable manufacturers. The scandal was not exposed by the official regulators but by an 'extended peer', an independent testing agency. Although sophisticated mathematical methods were not deployed in either of those fateful investigations, the two issues of competence and integrity are closely related. If so much of policy-relevant mathematical science is shot through with confusion and error, then the societal protections against malfeasance in any area are severely weakened.

Reflections

Those who expose and analyse crises like these are at risk of being considered as fomenters of cynicism and despair. But they are necessary for breaking the barriers of inertia, complacency and protectionism, which inhibit those constructive responses and reforms which, however difficult and painful, will be essential to the survival of the activity in a recognisable form. Those measures will arise in large part internally, from practitioners who want to do an honest and competent job; and there have already been many initiatives for restoring research integrity in one aspect or another. These efforts will benefit from complementary initiatives promoted by people and institutions outside the traditionally defined world of science. Citizens have a right to engage in this work, since misapplied advice and inappropriate regulatory measures adversely affect citizens in their ordinary lives. Wherever there are hazards and pollution created by science-based industry, the 'science of *bads'* is deployed, with structures of resources and incentives, and criteria of quality, that are like mirror-opposites to those of the sciences of the discovery of knowledge and the production of goods. In such situations, the extended peer communities of citizens can be crucial in creating a dialogue. They can open up issues for discussion and action, or scrutinise methodologies with the questions 'what-if?', 'what-about?' and 'why?', or even exhume uncomfortable knowledge. Such critical inquiry happens routinely within the

research communities in healthy fields; because of the special character of the science of *bads* of policy relevant sciences, these external stimulants have an important role to play.

Returning to the theme of the workshop, if there is indeed a prevalence of insignificant digits in the quantitative policy related sciences, then the meaning of this less than perfect state of affairs will be a matter of legitimate concern for citizens. Because of the highly technical nature of the material, their contribution will be less broadly based than in the case of pollution studies. The involved citizens may well include highly trained specialists, who find themselves independent of the interests that are promoting or protecting an objectionable situation. This already happens in disputes over planning or environmental policy, and in some places the processes are highly institutionalised, with 'stakeholders' given formal recognition and even government support. The assumption that 'scientists' (defined by qualifications and employment) have a monopoly of relevant scientific knowledge now belongs to a bygone age.

With an appreciation of this broader context, we can understand the significance of the JRC's promotion of the two workshops in June. The later workshop focused on an issue which is very likely to become salient in the near future: the challenge of quality in DIY science. We chose the term 'DIY' rather than 'citizen' in order to emphasise the mode of production, rather than the social location, of this sort of activity. The technology of production of scientific knowledge is being transformed, really as a new industrial revolution but occurring at breakneck speed compared to the original. Although still small and marginal, it is already affecting policy processes and mainstream science and technology in a variety of ways. The challenge is easily understood, in the light of the discussions at the 'significant digits' workshop: if mainstream science, with all its traditions and institutions of quality assurance, is itself in a crisis of quality, what hope is there that DIY science, essentially anarchic, will solve those same problems in the absence of time for learning and of institutions for social control? There is no easy answer, of course; but we believe that even to pose the problem for public discussion is important for creating the conditions under which it can be managed.

All this is part of a very broad movement for bringing science into democratic society, which now enjoys support from leading institutions in Europe and elsewhere. There will be huge problems to be solved, and doubtless a full ration of errors and failures. But out of this process there may well be created a new sort of scientific practice, with renewed vigour, creativity and morale. It is too early to predict its form, and how the new tasks of governance will be accomplished. But the present crisis can indeed become a turning point, analogous to the original scientific revolution out of which so much of the modern world came to be. In organising these two workshops of reflection on the challenges of the present and future, the JRC has made a valuable contribution to the process.

At the JRC

Many problems of the significance of digits, craft skills with numbers, deliberate and naïf misuse of numbers, etc. have been subject of different expert workshops at the JRC. The increasing scrutiny from extended peer communities who contribute extended facts and others way of knowing have also been reviewed. The JRC could continue its job of knowledge quality assurance not only through effecting in-house practices that take stock of these defects and offer remedies, but also to engage in training. It has the responsibility for such activity, as it is a boundary institution whose spheres of operation are at the intersection of science and policy. The analysis of bad practices is useful but not enough, and

it may indeed eventually become counterproductive. Training at different levels of operation at the European Commission (scientific and policy officers alike) is an option to be considered. We suggest that an extensive embedding of the users and trainers of the JRC reflexivity tools in its regular work would provide a great service to the European policies it supports, and also to the policy process in Europe and worldwide. This work could then be extended outside, to groups variously known as extended peer community, stakeholders, citizens or DIY scientists. Nothing would make a greater direct contribution to the health and vitality of the emerging peoples' science, than such an extended educational programme in awareness and quality.

References and Bibliography

Lakoff, G. (2004). Don't think of an elephant: know your values and frame the debate. London: Chelsea Green Publishing.

Lyotard, J.-F. (1979). La Condition postmoderne. Rapport sur le savoir, Paris: Minuit.

Mirowski, P. (2011). Science-Mart: Privatizing American Science, Cambridge MA: Harvard University Press.

Mirowski, P. (2013). Never Let a Serious Crisis Go to Waste: How Neoliberalism Survived the Financial Meltdown, Brooklyn: Verso Books.

Porter, T. M. (1995). Trust in Numbers, The Pursuit of Objectivity in Science and Public Life, Princeton.

Funtowicz, S. (2006). What is Knowledge Assessment? In Guimarães Pereira, Â., Guedes Vaz, S. and Tognetti, S. (eds) *Interfaces between Science and Society,* Sheffield: Greenleaf Publishers.

Ravetz, J. (1971). Scientific Knowledge and its Social Problems, Oxford University Press.

Ravetz, J. R. (1986). Usable knowledge, usable ignorance: Incomplete science with policy implications. In W. C. Clark & R. Munn (Eds.), *Sustainable development of the biosphere*. New York: IIASA/Cambridge University Press.

Urban, M. C. (2015). Accelerating extinction risk from climate change. *Science*, *348*(6234), 571-573.

Appendix

Abstracts of the contributions to the workshop Significant Digits: Responsible Use of quantitative Information held in Brussels on 9-10 June 2015[5].

Abstracts

Significant Digits: Opening

Vladimír Šucha

European Commission, Director General of DG Joint Research Centre, Brussels, BE

Director General Vladimír Šucha will introduce the workshop, outlining the relevance of this endeavour in the broader context of the DG Joint Research Centre's work and its commitment to quality scientific advice to policy, *viz à viz* and with special relevance to the new regulations and advice frameworks recently proposed by the European Commission.

The Significance of Digits

Jerome Ravetz

Institute for Science, Innovation and Society, Univ. of Oxford, UK

Andrea Saltelli and Mario Giampietro[6] have shown us how to solve the crisis of 'evidence-based' policy, through an enriched conception of quality of science. Here I would like to use my historical perspective to suggest how the problem came to be. Its origins lie in a certain conception of knowledge, which can be summed up as a faith in digits as nuggets of truth. Since this is so deep in our modern scientific culture as to pass unnoticed, its exposure and correction will not be a quick or straightforward process. But we must start somewhere, and here we are.

The existence of a crisis is beyond doubt. Economics, the king of the sciences of society, has been exposed as vacuous in its main function, namely explaining and helping to guide the running of the economy. And economics has long prided itself on being the physics of society. In this it has ignored the actual state of physics for

[5] The full agenda is available from: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC96441/kjna27387enn.pdf

[6] See The fallacy of evidence based policy – Available at: http://www.andreasaltelli.eu/file/repository/PaperDraftPolicyCartesianDream_16c.pdf

the last century, riven by uncertainty and paradox. For economists, and all the practitioners of social would-be technologies who have imitated it, the faith was simple. Given the reliance on numerical data and mathematical methods, what could possibly go wrong? The answer turned out to be 'everything'.

This delusionary faith has complex roots, in traditions of philosophical thought and social practice. Here I focus on digits, as a core element and symptom of the pathology. Other colleagues here have analysed the misuse and abuse of numerical information at great depth. For brevity would like to use just two examples of the unnoticed, or rather suppressed, contradictions in numerical information.

One is a variant on the classic 'fossils joke', where we consider the calculation: 65,000,000 – 3 = 64,999,997. This illustrates the ambiguity in the zero, functioning as either counter or filler, and its meaning depending on context. Thus even in digital information semantics sometimes dominates over syntax. Who knew that? The other is the question, how many significant digits should we use in expressing an 'error-bar'? Is it really meaningful to say that we know that (say) the 95% upper limit of probability of an estimate is 3.65 and not 3.64 or 3.66? If not, is there a clear meaning there at all? My point is that a practice that depends on the concealment of its confusions and contradictions will be particularly vulnerable to misuse and abuse. I opened a discussion of these issues in the chapter on 'Obscurities at the Foundations of Theoretical Science' in my earlier book.

I offer two historical parallels. Descartes' classic denunciation of humanist teaching on ethics, as "towering and magnificent palaces with no better foundation than sand and mud" might now become applied to the mathematical policy sciences. And the understanding and practice of scripture-based religion in the West was transformed in modern times by the critical study of its sources. Will this present crisis provide the opportunity for science to reflect and catch up?

Since so many policy issues now involve quantities with 'not even one significant digit', we need an appropriate new arithmetical language, based on 'soft numbers' using 'sparse digits' and a dynamical graphical arithmetic of uncertainty and quality.

On the extinction of craft skills with numbers: the case of "Overall, 7.9% of species are predicted to become extinct from climate change"

Jeroen Van Der Sluijs

Centre for the Study of the Sciences & the Humanities, University of Bergen, Norway

Since the establishment in the 1980s of science-policy interfaces around anthropogenic climate change attempts to quantify climate risks have produced various “magic numbers”. The classic example is the 1.5-4.5 °C range for the Earth’s climate sensitivity. Such numbers are produced in a particular way and within a particular context and are conditioned on a complex set of assumptions covering a wide range of scientific statuses ranging from crude speculation to well established knowledge. Once thrown over the disciplinary fence, important caveats tend to be ignored, uncertainties compressed and numbers used at face value. Poor practice in communication of uncertainty and the loss of what Jerome Ravetz calls “craft skills with numbers” has created a host of misunderstandings and miscommunication in quantitative information on climate change at the science policy and science society interfaces. This paper analyses the case of quantification of the risk that climate change poses to biodiversity. In 2004, Thomas *et al.* (doi:10.1038/nature02121) were the first to quantify extinction risks from climate change. On the basis of a highly aggregated species-area relationship model and climate projections of habitat loss, they predicted that by 2050 “15–37% of species in our sample of regions and taxa will be ‘committed to extinction’.”. Recently, based on a meta-analysis, Urban (2015 doi:10.1126/science.aaa4984) concluded that “overall, 7.9% of species are predicted to become extinct from climate change”. This paper will critically reflect on the meaning of the number “7.9%” and discuss the two papers from the viewpoint of craft skills with numbers and good practice in uncertainty communication.

Quantitative story telling as a therapy for hypocognition

Mario Giampietro

Institut de Ciència i Tecnologia Ambientals (ICTA) -Universitat Autonoma de Barcelona, Spain

Numbers in isolation do not carry meaning; they have to be always contextualised (examples of blunders and problematic formalizations using indicators). Numbers do not carry enough information for generating a robust integrated assessment - you

need vectors and matrices (data arrays) and grammars in order to transfer information across levels and dimensions. Especially important is to be aware of the distinction between information referring to types and information referring to special instances.

When dealing with the analysis of complex systems it is impossible to adopt a contextualisation "one size fits all" so, the more we try to formalise complex problems, the more we generate *hypocognition*.

The way out is an integration of different quantitative types of analysis properly contextualized that have to be handled simultaneously using semantic relations. But if one decides to adopt this solution one has to acknowledge the arbitrariness of the choice of your stories. The talk concludes with examples of quantitative story-telling information systems.

Fisheries quota advice for management: Significant scripts and significant digits

Dorothy J. Dankel

Centre for the Study of the Sciences & the Humanities, University of Bergen, Norway

The oldest and most prominent of scientific institutions in Europe with a mandate to provide advice on ecosystem-based management of the Ocean is the International Council for the Exploration of the Sea (ICES). ICES is made up of over 1000 scientists who are active in annual Expert Groups and produce annual reports with annual advice delivered as the “Total Allowable Catch” (TAC). The well-oiled TAC-machine does its job as dictated by ICES’ clients, but here I problematize the perception that TAC advice should be given as a single number. Often, a single and conclusive scientific answer will never be available for complex systems such as fisheries and marine ecosystems. In such cases, more research does not lead to less uncertainty, but can lead instead to unforeseen complexities (Van der Sluijs et al., 2005a, 2005b, 2010). Values are in dispute when the potential impacts of decisions based on uncertain models have very large biological and/or social consequences. Among ICES’ goals in its new Strategic Plan (2014-2018) is increased transparency and better and increased dialogue with stakeholders, however we still observe stakeholder reluctance to develop alternative ways of delivering advice. Furthermore, I use the term “scripts” as a reference to Goffmann (Giddens, 1998) to describe the changing roles of ICES scientists the last decade. We observe a shift from passive scientific reviews to active science that increasingly is in dialogue with stakeholders in the development of new management strategies. However, the plurality of roles of

single scientists is not arbitrary, and I argue that scripts and digits are inevitably intertwined. I conclude by introducing a method, the "Confidence Level Harvest Control Rule" for fisheries advice that potentially remedies both the script-juggling problem of scientists and the significant digit/significant model problem for society.

"Roles specify generalized expectations to which an individual has more or less closely to conform when in a particular situated context. […]. The self consists in an awareness of identity which simultaneously transcends specific roles and provides an integrating means of relating them to personal biography: and a set of dispositions for managing the transactions between motives and the expectations 'scripted' by particular roles"

(Giddens, 1988) [p. 258-259).

References:

Giddens, A. 1988. *Goffman as a systematic social theorist*. *In* Erving Goffman: Exploring the interaction order, pp. 250-279. Ed. by P. Drew, and A. Wootton. Polity Press.

Van der Sluijs, J. P. 2005a. Uncertainty as a monster in the science policy interface: four coping strategies. Water Science and Technology, 52: 87 – 92.

Van der Sluijs J.P.; Craye, M.; Funtowicz, S.; Kloprogge, P.; Ravetz, J.; and Risbey, J. 2005b. Combining Quantitative and Qualitative Measures of Uncertainty in Model based Environmental Assessment: the NUSAP System. Risk Analysis 25 (2) 481-492.

Van der Sluijs, J. P., van Est, R., and Riphagen, M. 2010. Beyond con-sensus: reflections from a democratic perspective on the interaction between climate politics and science. Current Opinion in Environmental Sustainability, 2: 409 – 415.

Pay no attention to the model behind the curtain

Philip B. Stark

Univ. of California, Berkeley, USA

Watch me pull a probability out of my model ... Presto! Typical attempts to quantify risk for policy makers involve inventing a stochastic model for a phenomenon; fitting some parameters in that model to data; then declaring that features of the fitted model, called "probabilities" within the model, magically apply to the real world. Pulling this probability rabbit from the analyst's hat generally involves several statistical and philosophical sleights of hand: confusing the map (the model) with the

territory (the phenomenon), confusing rates with probabilities, and distracting attention from the moment that probability entered the hat (i.e., the moment the stochastic model was assumed to have generated the data). Bedazzling the onlookers with a sparkly array of Greek symbols, heroic high-performance computing, and superficial attempts to quantify the uncertainty renders the show all the more dramatic.

The simplification of complexity: challenges of sustainability science for governance

Zora Kovavic

Institut de Ciència i Tecnologia Ambientals (ICTA) - Universitat Autònoma de Barcelona, Spain

The criticism to evidence based policy is a criticism to the idea that science can provide policy makers with all the information that is needed in order to decide for the common good (Funtowicz and Strand 2007). The assumption behind this model of science-policy is that (1) scientific information is a faithful representation of reality, whose interpretation and use are unequivocal, and that (2) policy decisions are based on the scientific information provided by the scientific community.

The first assumption is addressed by looking at instances of scientific information used to discuss policies. Based on the case of the indicators produced following the Beyond GDP Conference of 2007 and of the scientific evidence used in water management in Israel, I will give examples of the high level of ambiguity associated with scientific information and of the multiple representations of the same problem that can be produced by using different scales of analysis, different narratives and different time frames. The plurality of representations and perspectives that can be found in science suggest that scientific information reflects the normative stand of the analyst in relation to what is to be considered relevant in the framing of a problem – rather than a faithful representation of reality.

The second assumption is addressed by looking at the management of the financial crisis of 2008. In this case, policies seem to be irresponsive to the improvement of models, to more refined information, and to more data. In this situation, to what extent is it worth producing more accurate measurements, more quantitative information, and more indicators?

I argue that a better understanding of complexity can provide some of the tools needed to deal with uncertainty and pluralism. Complexity is defined as a systematic study of the way in which different perspectives are expressed in different pre-

analytical choices and in the resulting non-equivalent representations of the same problem. Complexity thus offers a way to assess the usefulness of the scientific information used in different policy contexts.

References:

Funtowicz, S.O. and Strand, R. 2007. Models of science and policy. In: Traavik, T. and Lim, L.C. (eds.). 2007. *Biosafety first: Holistic approaches to risk and uncertainty in genetic engineering and genetically modified organisms.* Trondheim: Tapir Academic Press.

Evidence based policy: handle with care

Andrea Saltelli

Centre for the Study of the Sciences & the Humanities, University of Bergen, Norway

The use of science for policy is at the core of a perfect storm generated by the insurgence of several concurrent crises: of science, of trust, of sustainability. The prevailing modern positivistic model of science for policy, known as 'evidence based policy', is based on dramatic simplifications and compressions of available perceptions of the state of affairs and possible explanations (*hypocognition*). Therefore this model can result in seriously flawed prescriptions.

The primacy of science to adjudicate political issues must pass through a serious assessment of the level of maturity and effectiveness of the various disciplines. The solution implies abandoning dreams of prediction, control and optimization obtained by relying on a limited set of simplified narratives to define a problem to be dealt with and move instead to an open exploration of a broader set of plausible and relevant stories.

We make examples of instrumental or otherwise vacuous use of evidence for policy. Computing climate's dollars are a case in point. We mention some strategies to spot problems and to tackle them.

Knowing what we don't know

John Kay

London School of Economics, UK

I review three model types – the cost benefit model (WEBTAG) used for assessing UK transport projects, the value for money models used to justify PFI (private finance initiative) schemes, and the value at risk (VAR) models widely employed in the financial sector. These have a common structure:

- write down the calculations you would make if the world were completely known
- since very little of it is in fact known, almost all the numbers in the cells of the spreadsheet are invented
- a standard template is applied to a widely varied class of problems.

I will analyse the deficiencies of these approaches and what might be done instead.

KJ-04-15970-EN-N

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation

Publications Office

doi: 10.2760/9793

ISBN 978-92-79-54261-9